

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,746

DATE: 05/22/2001

TIME: 15:24:25

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05222001\I840746.raw

```

2 <110> APPLICANT: Chen, Huei-Mei
3     Honchell, Cynthia D.
4     Tang, Y. Tom
6 <120> TITLE OF INVENTION: Mucin-Related Tumor Marker
8 <130> FILE REFERENCE: PC-0039 US
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/840,746
C--> 11 <141> CURRENT FILING DATE: 2001-04-23
13 <160> NUMBER OF SEQ ID NOS: 20
14 <170> SOFTWARE: PERL Program
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 946
18 <212> TYPE: PRT
19 <213> ORGANISM: Homo sapiens
21 <220> FEATURE:
22 <221> NAME/KEY: misc_feature
23 <223> OTHER INFORMATION: Incyte ID No: 182514CD1
25 <400> SEQUENCE: 1
26   Met Ser Gln Thr Glu Thr Val Ser Arg Ser Val Ala Pro Met Arg
27       1           5           10           15
28   Gly Gly Glu Ile Thr Ala His Trp Leu Leu Thr Asn Ser Thr Thr
29           20           25           30
30   Ser Ala Asp Val Thr Gly Ser Ser Ala Ser Tyr Pro Glu Gly Val
31           35           40           45
32   Asn Ala Ser Val Leu Thr Gln Phe Ser Asp Ser Thr Val Gln Ser
33           50           55           60
34   Gly Gly Ser His Thr Ala Leu Gly Asp Arg Ser Tyr Ser Glu Ser
35           65           70           75
36   Ser Ser Thr Ser Ser Ser Glu Ser Leu Asn Ser Ser Ala Pro Arg
37           80           85           90
38   Gly Glu Arg Ser Ile Ala Gly Ile Ser Tyr Gly Gln Val Arg Gly
39           95          100          105
40   Thr Ala Ile Glu Gln Arg Thr Ser Ser Asp His Thr Asp His Thr
41          110          115          120
42   Tyr Leu Ser Ser Thr Phe Thr Lys Gly Glu Arg Ala Leu Leu Ser
43          125          130          135
44   Ile Thr Asp Asn Ser Ser Ser Ser Asp Ile Val Glu Ser Ser Thr
45          140          145          150
46   Ser Tyr Ile Lys Ile Ser Asn Ser Ser His Ser Glu Tyr Ser Ser
47          155          160          165
48   Phe Ser His Ala Gln Thr Glu Arg Ser Asn Ile Ser Ser Tyr Asp
49          170          175          180
50   Gly Glu Tyr Ala Gln Pro Ser Thr Glu Ser Pro Val Leu His Thr
51          185          190          195
52   Ser Asn Leu Pro Ser Tyr Thr Pro Thr Ile Asn Met Pro Asn Thr
53          200          205          210
54   Ser Val Val Leu Asp Thr Asp Ala Glu Phe Val Ser Asp Ser Ser
55          215          220          225

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,746

DATE: 05/22/2001

TIME: 15:24:25

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05222001\I840746.raw

56	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Gly	Pro	Pro
57					230					235					240
58	Leu	Pro	Leu	Pro	Ser	Val	Ser	Gln	Ser	His	Leu	Phe	Ser	Ser	
59					245					250					255
60	Ile	Leu	Pro	Ser	Thr	Arg	Ala	Ser	Val	His	Leu	Leu	Lys	Ser	Thr
61					260					265					270
62	Ser	Asp	Ala	Ser	Thr	Pro	Trp	Ser	Ser	Ser	Pro	Ser	Pro	Leu	Pro
63					275					280					285
64	Val	Ser	Leu	Thr	Thr	Ser	Thr	Ser	Ala	Pro	Leu	Ser	Val	Ser	Gln
65					290					295					300
66	Thr	Thr	Leu	Pro	Gln	Ser	Ser	Ser	Thr	Pro	Val	Leu	Pro	Arg	Ala
67					305					310					315
68	Arg	Glu	Thr	Pro	Val	Thr	Ser	Phe	Gln	Thr	Ser	Thr	Met	Thr	Ser
69					320					325					330
70	Phe	Met	Thr	Met	Leu	His	Ser	Ser	Gln	Thr	Ala	Asp	Leu	Lys	Ser
71					335					340					345
72	Gln	Ser	Thr	Pro	His	Gln	Glu	Lys	Val	Ile	Thr	Glu	Ser	Lys	Ser
73					350					355					360
74	Pro	Ser	Leu	Val	Ser	Leu	Pro	Thr	Glu	Ser	Thr	Lys	Ala	Val	Thr
75					365					370					375
76	Thr	Asn	Ser	Pro	Leu	Pro	Pro	Ser	Leu	Thr	Glu	Ser	Ser	Thr	Glu
77					380					385					390
78	Gln	Thr	Leu	Pro	Ala	Thr	Ser	Thr	Asn	Leu	Ala	Gln	Met	Ser	Pro
79					395					400					405
80	Thr	Phe	Thr	Thr	Thr	Ile	Leu	Lys	Thr	Ser	Gln	Pro	Leu	Met	Thr
81					410					415					420
82	Thr	Pro	Gly	Thr	Leu	Ser	Ser	Thr	Ala	Ser	Leu	Val	Thr	Gly	Pro
83					425					430					435
84	Ile	Ala	Val	Gln	Thr	Thr	Ala	Gly	Lys	Gln	Leu	Ser	Leu	Thr	His
85					440					445					450
86	Pro	Glu	Ile	Leu	Val	Pro	Gln	Ile	Ser	Thr	Glu	Gly	Gly	Ile	Ser
87					455					460					465
88	Thr	Glu	Arg	Asn	Arg	Val	Ile	Val	Asp	Ala	Thr	Thr	Gly	Leu	Ile
89					470					475					480
90	Pro	Leu	Thr	Ser	Val	Pro	Thr	Ser	Ala	Lys	Glu	Met	Thr	Thr	Lys
91					485					490					495
92	Leu	Gly	Val	Thr	Ala	Glu	Tyr	Ser	Pro	Ala	Ser	Arg	Ser	Leu	Gly
93					500					505					510
94	Thr	Ser	Pro	Ser	Pro	Gln	Thr	Thr	Val	Val	Ser	Thr	Ala	Glu	Asp
95					515					520					525
96	Leu	Ala	Pro	Lys	Ser	Ala	Thr	Phe	Ala	Val	Gln	Ser	Ser	Thr	Gln
97					530					535					540
98	Ser	Pro	Thr	Thr	Leu	Ser	Ser	Ser	Ala	Ser	Val	Asn	Ser	Cys	Ala
99					545					550					555
100	Val	Asn	Pro	Cys	Leu	His	Asn	Gly	Glu	Cys	Val	Ala	Asp	Asn	Thr
101					560					565					570
102	Ser	Arg	Gly	Tyr	His	Cys	Arg	Cys	Pro	Pro	Ser	Trp	Gln	Gly	Asp
103					575					580					585
104	Asp	Cys	Ser	Val	Asp	Val	Asn	Glu	Cys	Leu	Ser	Asn	Pro	Cys	Pro

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,746

DATE: 05/22/2001

TIME: 15:24:25

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05222001\I840746.raw

105		590		595		600
106	Ser Thr Ala Thr	Cys Asn Asn Thr Gln	Gly Ser Phe Ile Cys	Lys		
107		605		610		615
108	Cys Pro Val Gly	Tyr Gln Leu Glu Lys	Gly Ile Cys Asn Leu	Val		
109		620		625		630
110	Arg Thr Phe Val	Thr Glu Phe Lys Leu	Lys Arg Thr Phe Leu	Asn		
111		635		640		645
112	Thr Thr Val Glu	Lys His Ser Asp Leu	Gln Glu Val Glu Asn	Glu		
113		650		655		660
114	Ile Thr Lys Thr	Leu Asn Met Cys Phe	Ser Ala Leu Pro Ser	Tyr		
115		665		670		675
116	Ile Arg Ser Thr	Val His Ala Ser Arg	Glu Ser Asn Ala Val	Val		
117		680		685		690
118	Ile Ser Leu Gln	Thr Thr Phe Ser Leu	Ala Ser Asn Val Thr	Leu		
119		695		700		705
120	Phe Asp Leu Ala	Asp Arg Met Gln Lys	Cys Val Asn Ser Cys	Lys		
121		710		715		720
122	Ser Ser Ala Glu	Val Cys Gln Leu Leu	Gly Ser Gln Arg Arg	Ile		
123		725		730		735
124	Phe Arg Ala Gly	Ser Leu Cys Lys Arg	Lys Ser Pro Glu Cys	Asp		
125		740		745		750
126	Lys Asp Thr Ser	Ile Cys Thr Asp Leu	Asp Gly Val Ala Leu	Cys		
127		755		760		765
128	Gln Cys Lys Ser	Gly Tyr Phe Gln Phe	Asn Lys Met Asp His	Ser		
129		770		775		780
130	Cys Arg Ala Cys	Glu Asp Gly Tyr Arg	Leu Glu Asn Glu Thr	Cys		
131		785		790		795
132	Met Ser Cys Pro	Phe Gly Leu Gly Gly	Leu Asn Cys Gly Asn	Pro		
133		800		805		810
134	Tyr Gln Leu Ile	Thr Val Val Ile Ala	Ala Ala Gly Gly Gly	Leu		
135		815		820		825
136	Leu Leu Ile Leu	Gly Ile Ala Leu Ile	Val Thr Cys Cys Arg	Lys		
137		830		835		840
138	Asn Lys Asn Asp	Ile Ser Lys Leu Ile	Phe Lys Ser Gly Asp	Phe		
139		845		850		855
140	Gln Met Ser Pro	Tyr Ala Glu Tyr Pro	Lys Asn Pro Arg Ser	Gln		
141		860		865		870
142	Glu Trp Gly Arg	Glu Ala Ile Glu Met	His Glu Asn Gly Ser	Thr		
143		875		880		885
144	Lys Asn Leu Leu	Gln Met Thr Asp Val	Tyr Tyr Ser Pro Thr	Ser		
145		890		895		900
146	Val Arg Asn Pro	Glu Leu Glu Arg Asn	Gly Leu Tyr Pro Ala	Tyr		
147		905		910		915
148	Thr Gly Leu Pro	Gly Ser Arg His Ser	Cys Ile Phe Pro Gly	Gln		
149		920		925		930
150	Tyr Asn Pro Ser	Phe Ile Ser Asp Glu	Ser Arg Arg Arg Asp	Tyr		
151		935		940		945
152	Phe					
155	<210> SEQ ID NO: 2					

## RAW SEQUENCE LISTING

DATE: 05/22/2001

PATENT APPLICATION: US/09/840,746

TIME: 15:24:25

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05222001\I840746.raw

156 &lt;211&gt; LENGTH: 6952

157 &lt;212&gt; TYPE: DNA

158 &lt;213&gt; ORGANISM: Homo sapiens

160 &lt;220&gt; FEATURE:

161 &lt;221&gt; NAME/KEY: misc\_feature

162 &lt;223&gt; OTHER INFORMATION: Incyte ID No: 182514CB1

164 &lt;400&gt; SEQUENCE: 2

```

165   gttc gatgaa agaattgccg cttttcaaac aaagagtggg acagcctcgg agatggggaac 60
166   agagagggcg atggggctgt cagaagaatg gactgtgcac agccaagagg ccaccacttc 120
167   ggcttggagc ccttcctttc ttctgtcttt ggagatggga gagctgacca cgccttctag 180
168   gaagagaaat tcctcaggac cagatctctc ctggctgcat ttctacagga cagcagcttc 240
169   ctctcctctc ttagaccttt cctcaccttc tgaaagtaca gagaagctta acaactccac 300
170   tggcctccag agctcctcag tcagtcaaac aaagacaatg catgttgcta ccgtgttcac 360
171   tgatggtggc ccgagaacgc tgcgatcttt gacggtcagt ctgggacctg tgagcaagac 420
172   agaaggcttc cccaaggact ccagaattgc cagacttca tcctcagtcc ttctttcacc 480
173   ctctgcagtg gaatcgagaa gaaacagtag agtaactggg aatccagggg atgaggaatt 540
174   cattgaacca tccacagaaa atgaatttgg acttacgtct ttgctgggca aaatgattcc 600
175   ccaacctttg gagaacatca gcttgccagc agctctgagg tgcaaaatgg aagtcccacg 660
176   tctcagactg agactgtgtc taggtcagtc gcacccatga gaggtggaga gatcactgca 720
177   cactggctct tgaccaacag cacaacatct gcagatgtga caggaagctc tgcttcatat 780
178   cctgaagggtg tgaatgtctc agtgttgacc cagttctcag actctactgt acagtctgga 840
179   ggaagtccaa cagcattggg agataggagt tattcagagt cttcatctac atcttcctcg 900
180   gaaagcttga attcatcagc accacgttga gaacgttcaa tcgctgggat tagctacggt 960
181   caagtgcgtg gcacagctat tgaacaaaagg acttccagcg accacacaga ccacacctac1020
182   ctgtcatcta ctttcaccaa aggagaacgg gcgttactgt ccattacaga taacagttcal1080
183   tcctcagaca ttgtggagag ctcaacttct tatattaaaa tctcaaactc ttcacattcal1140
184   gagtattcct ctttttctca tgctcagact gagagaagta acatctcatc ctatgacggg1200
185   gaatatgtct agccttctac tgagtcgcca gttctgcata catccaacct tccgtcctac1260
186   acaccaccca ttaatatgcc gaacacttcg gttgttctgg aactgatgc tgagtttgtt1320
187   agtgactcct cctcctctc ttctcctcc tcctcttctt cttcttcagg gcctcctttg1380
188   cctctgccct ctgtgtcaca atcccacat ttattttcat caattttacc atcaaccagg1440
189   gcctctgtgc atctactaaa gtctacctct gatgcatcca caccatggtc ttctcacca1500
190   tcacctttac cagtatcctt aacgacatct acatctgccc cactttctgt ctcaaaaacal1560
191   accttgccac agtcatcttc taccctgtc ctgcccaggg caagggagac tcctgtgact1620
192   tcatttcaga catcaacaat gacatcattc atgacaatgc tccatagtag taaaactgcal1680
193   gaccttaaga gccagagcac cccacaccaa gagaaagtca ttacagaatc aaagtcaccal1740
194   agcctggtgt ctctgccac agagtccacc aaagctgtaa caacaaactc tcctttgcct1800
195   ccacctttaa cagagtcctc cacagagcaa acccttccag ccacaagcac caacttagcal1860
196   caaatgtctc caactttcac aactaccatt ctgaagacct ctgacctct tatgaccact1920
197   cctggcacc tgtcaagcac agcatctctg gtcactggcc ctatagccgt acagactaca1980
198   gctggaaaac agctctcgct gacccatcct gaaatactag ttctcaaat ctcaacagaa2040
199   ggtggcatca gcacagaaag gaaccgagt attgtggatg ctaccactgg attgatccct2100
200   ttgaccagtg taccacatc agcaaaagaa atgaccacaa agcttggcgt tacagcagag2160
201   tacagcccag cttcacgttc cctcggaaac tctccttctc cccaaaccac agttgtttcc2220
202   acggctgaag acttggctcc caaatctgcc acctttgctg ttcagagcag cacacagtca2280
203   ccaacaacac tgcctcttc agcctcagtc aacagctgtg ctgtgaacct ttgtcttcac2340
204   aatggcgaat gcgtcgaga caacaccagc cgtggctacc actgcaggtg cccgccttcc2400
205   tggcaagggg atgattgcag tgtggatgtg aatgagtgcc tgtcgaacct ctgcccaccc2460
206   acagccacgt gcaacaatac tcagggatcc tttatctgca aatgcccggg tgggtaccag2520

```

## RAW SEQUENCE LISTING

DATE: 05/22/2001

PATENT APPLICATION: US/09/840,746

TIME: 15:24:25

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05222001\I840746.raw

```

207 ttggaaaaaag ggatatgcaa tttgggttaga accttcgtga cagagtttaa attaaagaga2580
208 actttttctta atacaactgt ggaaaaacat tcagacctac aagaagttga aaatgagatc2640
209 accaaaacgt taaatatgtg tttttcagcg ttacctagt acatccgatc tacagttcac2700
210 gcctctaggg agtccaacgc ggtggtgatc tcaactgcaa caaccttttc cctggcctcc2760
211 aatgtgacgc tatttgacct ggctgatagg atgcagaaat gtgtcaactc ctgcaagtcc2820
212 tctgctgagg tctgccagct cttgggatct cagaggcgga tcttttagagc gggcagcttg2880
213 tgcaagcgga agagtcccga atgtgacaaa gacacctcca tctgcaactga cctggacggc2940
214 gttgccctgt gccagtgcaa gtcgggatac tttcagttca acaagatgga ccactcctgc3000
215 cgagcatgtg aagatggata taggcttgaa aatgaaacct gcatgagttg cccattttggc3060
216 cttggtggtc tcaactgtgg aaaccctat cagcttatca ctgtggtgat cgcagccgcg3120
217 ggaggtgggc tctgtctcat cctaggcatc gcaactgattg ttacctgttg cagaaagaat3180
218 aaaaatgaca taagcaaaact catcttcaaa agtggagatt tccaaatgtc cccatatgtc3240
219 gaatacccca aaaatcctcg ctcaacaaga tggggccgag aagctattga aatgcatgag3300
220 aatggaagta ccaaaaacct cctccagatg acggatgtgt actactcgcc tacaagtgt3360
221 aggaatccag aacttgaacg aaacggactc taccggcct acactggact gccaggatca3420
222 cggcatttctt gcattttccc cggacagtat aacctgtctt tcatcagtga tgaaagcaga3480
223 agaagagact acttttaagt ccaggagaga gagggactca ttgctctgag ccagtcacct3540
224 gggacctctg ctcaaggagc cgcaccagga ggctgcgccc aggattttgtc gggagccacg3600
225 ctgagtggca agcaggaaga gggacaggca tgcggggcgt gaccacagtg gaggagacag3660
226 gtggatgtgg aaccacaggc tgctcattca gcacctttgt tgttactgtg aacgtgaatg3720
227 tgggccagta tcaagagagt ctctctgagt gactgcacca tggcactggc accaggcgca3780
228 ctattagcca gggcagacca ctagacttca gtgcaggac ctggttttcc cttcgtttgc3840
229 acttttagtaa attgggtggg aggtttcctt ttggatctgt tttgagactg ttccagaaag3900
230 aaggcttcct tccccgagac acttccatag gcagcaattt ggtgattcat ttgcagcaaa3960
231 atactggctt gttaattatt ttctgcccga gcgctgcgt gctaaacaac agatgaggat4020
232 gagcgtaacca ctgaagtctg aagatgtcgc cattgaacgg acagtgtttt catatgtttc4080
233 taggttgtct tatgctacag tttccaagcc agccccaca gtgaggaaat gtgtgaggca4140
234 ccgcacacaa ctgcaatgtg ttttttaagt caaggtgaca catgtattta agattttttt4200
235 ttaaaatctc tttgcagtta aatctcactt tttcaaaaca gcctggatca gggcaaaaca4260
236 acttatattt ggttttagct ggaggctcag caggcagatt gcaggcaggg gggcactttt4320
237 catccatgag ggcccagcct ggggcctggg actctgatca ccattgtgga ggccagaggc4380
238 agctgcgtat ggaggagaaa tgtcaaaactg aacgcagggt tcaccactct aggaaagcag4440
239 cttgttgagc ccctgcagct ggatgtggtt agagggatgg gctgaatagg caggttagat4500
240 ttcctgcatac aacagtgcct tgggaagctg tgtggattcc tgaggaagaa caggagaccg4560
241 agatggagcc acacatgagt ttgctcaccg gctactgcag cactttgtac ccagaatctc4620
242 atgtccacaa accccatgta aactttcaac cactcaaagc tgtttattcg gctgaagaaa4680
243 taactttttt ttctcaccca gtcatttgta cctcttcata tggctgtgtc gcaccctcca4740
244 gaaacgtggt tatacttcca gtcagtgtgg gagaactgaa gacttccggt tggtcgagga4800
245 actgagggtt gaccttcggg aaggaagttc cactcatctt atttattatg cctgtgatgt4860
246 gggtcctgcc agggagacat ccagtaactg gtgtctttaa ttgccacctg gggaaactgt4920
247 ttatttgccc ttctttgggg catcctggtt ttggatgaag tgaggggaat acagaggtaa4980
248 aagaattgtc tccaccctga agcggggagt cccgcttcac atttctggaa atggtgcagc5040
249 cactggggac agttctgccc cgggcattgt tgtttcttca aggtcctcta aatataatcc5100
250 ctattcttac ataatccttg gccctgatgg ttttaagcaa gaactcctgt gtcccatggt5160
251 ctccaccact caccatcacc ctgctgtagc aagagtccta gtcaggggag gtgcatttta5220
252 gtagttaaat tgcacttatc catgagataa ataaaaggag aactgttttt atcagtgagg5280
253 gctaacctaa aatttcaaag tgtcgccttt ttgaaatctt gggcctctct ctctgtagaa5340
254 ccaatggccc tttgtggctc acggcctcgc acctaactgg agagttctga gctcctgcag5400
255 ctccactgag cccacagact aggtctcttg gctccttccg cagcatgcct gctcaccccc5460

```

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/840,746

DATE: 05/22/2001

TIME: 15:24:26

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05222001\I840746.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:558 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15

L:559 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15

L:578 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16

L:579 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16

L:580 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16

L:581 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16

L:600 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:602 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17